

# Kinetic models of collective decision-making in the presence of equality bias

Lorenzo Pareschi\*, Pierluigi Vellucci†, Mattia Zanella‡

June 21, 2016

## Abstract

We introduce and discuss kinetic models describing the influence of the competence in the evolution of decisions in a multi-agent system. The original exchange mechanism, which is based on the human tendency to compromise and change opinion through self-thinking, is here modified to include the role of the agents' competence. In particular, we take into account the agents' tendency to behave in the same way as if they were as good, or as bad, as their partner: the so-called *equality bias*. This occurred in a situation where a wide gap separated the competence of group members. We discuss the main properties of the kinetic models and numerically investigate some examples of collective decision under the influence of the equality bias. The results confirm that the equality bias leads the group to suboptimal decisions.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Modeling opinion and competence</b>	<b>3</b>
2.1	Evolution of competence . . . . .	3
2.2	The dynamics of competence based decisions . . . . .	4
2.3	Dynamics of decisions under equality bias . . . . .	6
<b>3</b>	<b>A Boltzmann model for opinion and competence</b>	<b>7</b>
3.1	Collective decision and variance . . . . .	8
<b>4</b>	<b>Fokker-Planck approximation</b>	<b>10</b>
4.1	Stationary states for the marginal density . . . . .	12
<b>5</b>	<b>Numerics</b>	<b>13</b>
5.1	Test 1: collective decision under equality bias . . . . .	13
5.2	Test 2: competence driven optimal decision . . . . .	14
5.3	Test 3: bounded confidence case . . . . .	15

\*University of Ferrara, Department of Mathematics and Computer Science, Via N. Machiavelli 35 44121 Ferrara, Italy ([lorenzo.pareschi@unife.it](mailto:lorenzo.pareschi@unife.it)).

†University of Rome 1, Department of Basic and Applied Sciences for Engineering, Via A. Scarpa 16 00161, Roma, Italy ([pierluigi.vellucci@sbai.uniroma1.it](mailto:pierluigi.vellucci@sbai.uniroma1.it)).

‡University of Ferrara, Department of Mathematics and Computer Science, Via N. Machiavelli 35 44121 Ferrara, Italy ([mattia.zanella@unife.it](mailto:mattia.zanella@unife.it)).

## 1 Introduction

The fallibility of human judgement is evident in our everyday life, especially regarding our self-evaluation ability. Several tests have been designed in cognitive psychology and clinical research in order to find an experimental evidence for this phenomenon, see [12, 13, 19, 20] and the references therein, showing that subjects are in general overconfident about the correctness of their belief. This lack in *metacognition*, i.e. the self-assessment of our own knowledge skills, goes hand in hand with the grade of competence of each subject.

The correlation between *competence* and *metacognitive skills* is somehow double and might be summarized in the following sentence: “the same knowledge that underlies the ability to produce correct judgement is also the knowledge that underlies the ability to recognize correct judgement” [22]. Here the authors found a systematic bias of the most incompetent agents on their metacognition than the most experts; behavior which is usually known as *Dunning-Kruger effect*. In other words, incompetence, is not only follows by poor choices but also disable to recognize that these are wrong or improvable. Furthermore the overconfidence of the novices emerges together with the underconfidence of highly competent individuals which tend to negatively estimate their skills.

This coupled deviation from an objective self-evaluation of personal abilities has been recently investigated in [24], where authors asked how people deal with individual differences in competence, in the context of a collective perceptual decision-making task, developing a metric for estimating how participants weight their partner’s opinion relative to their own. Empirical experiments, replicated across three slightly different countries as Denmark, Iran, and China, show how participants assigned nearly equal weights to each other’s opinions regardless of the real differences in their competence. This *equality bias*, whereby people behave as if they are as good or as bad as their partner, is particularly costly for a group when a competence gap separates its members.

Drawing inspiration from these recent results, we propose a multi-agent model which takes into account the influence of competence during the formation of a collective decision [7, 28]. After the seminal models for wealth/opinion exchange for a multi-agent system introduced in [8, 32] some recent works considered additional parameters to quantify the personal knowledge or conviction [7, 10, 11, 28] or constrained versions of these models [1, 2, 21]. For example, individuals with high conviction are resistant to change opinion, and can play the role of leaders [2, 11]. In [23] there exists a threshold conviction beyond of which one of the two choices provided to the individuals prevails, spontaneously breaking the existing symmetry of the initial set-up.

More precisely, we introduce a binary exchange mechanism for opinion and competence deriving a kinetic equation of Boltzmann-type [1, 2, 7, 28, 29, 32]. The binary collision terms for competence and opinion describe different processes:

- the competence evolution depends on a social background in which individuals grow and on the possibility for less competent agents to learn from the more competent ones;
- the opinion dynamics is based on a competence based compromise process including an equality bias effect and change of opinion through self-thinking;
- agents are driven toward an a-priori correct choice in dependence on their competence grade;

In order to derive a nonlinear equation of Boltzmann-type for the joint evolution of competence and opinion in the limit of a large number of interacting agents we resort to the principles of classical kinetic theory (we refer to the recent monograph [29] for an introduction to the subject). Furthermore, to simplify the study of the asymptotic behavior of the model, we obtain a Fokker-Planck approximation of the dynamic in the so-called quasi-invariant scaling.

The rest of the paper is organized as follows, in Section 2 we introduce the binary interaction model for competence and opinion. We discuss the competence-based interactions between agents formulating a definition of collective optimal decision which is coherent with the experimental setting of [24]. Then the equality bias function is introduced acting as a modification of the effective competence into perceived competence. In Section 3 we derive the Boltzmann-type model and study the evolution of the moments under some specific assumptions. The Fokker-Planck approximation is then obtained in Section 4, and we derive the stationary marginal density of the competence variable. Finally in Section 5 we present several numerical experiments which show that the model is capable to describe correctly the decision making process based on agents competence and to include the equality bias effects. The latter, as expected, drive the system towards suboptimal decions.

## 2 Modeling opinion and competence

In this section we discuss the modeling of opinion dynamics through binary exchanges, the analogous of dyadic interaction in the reference experimental literature [4, 24]. The mathematical approach follows several recent works on alignment processes in socio-economical dynamics [7, 11, 14, 25, 26, 28, 29, 32].

### 2.1 Evolution of competence

It is evident that one of the main factors playing a role is the social background in which an individual grows and lives, then it is natural to assume that competence is, in part, inherited from the environment. Moreover, we clearly have the possibility to learn a specific competence during interactions with other, more competent, agents.

Similarly to the works [7, 28, 29] we describe the evolution of competence of an individual in terms of a scalar parameter  $x \in X$  where  $X \subset \mathbb{R}^+$ . Let  $z \in \mathbb{R}^+$  be the degree of competence achieved from the background in each interaction; in what follows we will always suppose that  $C(z)$  is a distribution with bounded mean

$$\int_{\mathbb{R}^+} C(z)dz = 1, \quad \int_{\mathbb{R}^+} zC(z)dz = m_B. \quad (2.1)$$

We define the new amount of competence after a binary interaction of agents with competence  $x$  and  $x_*$  as follows

$$\begin{cases} x' &= (1 - \lambda(x))x + \lambda_C(x)x_* + \lambda_B(x)z + \kappa x \\ x'_* &= (1 - \lambda(x_*))x_* + \lambda_C(x_*)x + \lambda_B(x_*)z + \kappa x_*, \end{cases} \quad (2.2)$$

where  $\lambda(\cdot)$ ,  $\lambda_C(\cdot)$  and  $\lambda_B(\cdot)$  quantify the amounts of competence lost by individual by the natural process of forgetfulness, the competence gained thanks to the interaction with other agents and the expertise gained from the background respectively, while  $\kappa$  is a zero-mean random variable with finite second order moment  $\sigma_\kappa^2$ , taking into account the possible unpredictable changes of the competence process. A possible choice for  $\lambda_C(x)$  is  $\lambda_C(x) = \lambda_C \chi(x \geq \bar{x})$ , where  $\chi(\cdot)$  is the indicator function and  $\bar{x} \in X$  a minimum level of competence required to the agents' for increasing their own skills by interacting with the other agents.

With the dynamics (2.2) we introduced a general process in which agents respectively loose and gain competence interacting with the other agents and with the background. It is reasonable then to assume that both the processes of acquisition and loss, which are weighted by the coefficients  $\lambda$ ,  $\lambda_C$  and  $\lambda_B$ , are bounded by zero. Thus if  $\lambda \in [\lambda_-, \lambda_+]$ , with  $\lambda_- > 0$  and  $\lambda_+ < 1$ , and  $\lambda_C(x), \lambda_B(x) \in$

$[0, 1]$  the random part may be chosen to satisfy  $\kappa \geq -1 + \lambda_+$ . For example,  $\kappa$  may be uniformly distributed in  $[-1 + \lambda_+, 1 - \lambda_+]$ .

Let  $g(x, t)$  be the density function of individuals with competence  $x \in X \subset \mathbb{R}^+$  at time  $t \geq 0$ . Resorting to the standard Boltzmann-type setting, we refer to [29] for an extensive treatment, it is possible to describe in weak form the evolution of such density function as follows

$$\frac{d}{dt} \int_X \psi(x) g(x, t) dx = \left\langle \int_{\mathbb{R}^+ \times X} (\psi(x') - \psi(x)) g(x, t) C(z) dx dz \right\rangle, \quad (2.3)$$

where  $x'$  is the post-interaction competence given in (2.2), the brackets  $\langle \cdot \rangle$  indicate the expectation with respect to the random variable  $\kappa$  and  $\psi(\cdot)$  is a test function. By imposing  $\psi(x) = x$  we obtain an equation for the evolution of the mean-competence  $m_x(t)$

$$\begin{aligned} \frac{d}{dt} m_x(t) = \frac{1}{2} \left( \int_{X^2} (\lambda_C(x) - \lambda(x)) (x + x_*) g(x, t) g(x_*, t) dx dx_* + \right. \\ \left. \int_{\mathbb{R}^+ \times X} \lambda_B(x) z g(x, t) C(z) dx dz \right), \end{aligned} \quad (2.4)$$

which, for  $\lambda_C(x) = \lambda_C, \lambda_B(x) = \lambda_B, \lambda(x) = \lambda$ , yields

$$\frac{d}{dt} m_x(t) = -(\lambda - \lambda_C) m_x(t) + \lambda_B m_B, \quad (2.5)$$

whose solution is given by

$$m_x(t) = m_x(0) e^{-(\lambda - \lambda_C)t} + \frac{\lambda_B m_B}{\lambda - \lambda_C} (1 - e^{-(\lambda - \lambda_C)t}). \quad (2.6)$$

Therefore if  $\lambda > \lambda_C$  we obtain the asymptotic exponential convergence of the mean competence  $m_x$  toward  $\lambda_B m_B / (\lambda - \lambda_C)$ . Note that, if we assume  $\lambda = \lambda_B + \lambda_C$  we have that the average competence of the system tends to the mean competence induced by the variable  $z \in \mathbb{R}^+$ . Finally we remark that, compared to previous models where the notion of knowledge/conviction were introduced [7, 28], here we have a fully binary dynamic which includes also the possibility to increase the agent's competence as a result of the interaction with the other agents.

## 2.2 The dynamics of competence based decisions

Let us now consider a system of binary interacting agents, each of them endowed with two quantities  $(w, x)$  representing its opinion concerning a certain decision and competence respectively. Let  $I = [-1, 1]$  be the set of possible opinions for each interaction where the two extremal points  $\pm 1$  represent the two alternative decisions.

Two agents identified by the couples  $(x, w)$  and  $(x_*, w_*)$  modify their opinion after interaction according to the following rules

$$\begin{cases} w' = w - \alpha_S S(x)(w - w_d) - \alpha_P P(x, x_*; w, w_*)(w - w_*) + \eta D(x, w) \\ w'_* = w_* - \alpha_S S(x_*)(w_* - w_d) - \alpha_P P(x_*, x; w_*, w)(w_* - w) + \eta_* D(x_*, w_*), \end{cases} \quad (2.7)$$

where  $w, w_* \in I$  denote the pre-interaction opinions and  $w', w'_*$  the opinions after the exchange of information between the two agents. The positive function  $S(\cdot)$  drives the agent toward the correct choice  $w_d \in \{-1, 1\}$  with a rate dependent on its competence level, representing an individual decision making strength. For example, a possible choice for the function  $S(\cdot)$  is  $S(x) = \text{const.} > 0$  if  $x > \bar{x}$  and  $S(x) = 0$  elsewhere. In this case the agent needs to achieve a competence threshold

$\bar{x}$  in order to perceive the correct choice. Observe that in (2.7) we introduced an interaction function  $P(\cdot, \cdot; \cdot, \cdot)$  depending both on the competence and on the opinion of the pair of agents. The nonnegative parameters  $\alpha_S, \alpha_P$  characterize the drift toward the target opinion  $w_d \in I$  and the interaction rate, respectively. The random variables  $\eta, \eta_*$  are centered and with the same distribution  $\Theta$  with finite variance  $\sigma^2$ . The function  $D(\cdot, \cdot) \geq 0$  represents the local relevance of the diffusion for a given opinion and competence, whereas the evolution of the competences  $x, x_*$  are given by (2.2).

In absence of diffusion  $\eta, \eta_* = 0$  and for a constant drift  $S(\cdot) = S \leq 1$  we have

$$|w' - w'_*| \leq |1 - \alpha_S S - \alpha_P (P(x, x_*; w, w_*) + P(x_*, x; w_*, w))| |w - w_*|. \quad (2.8)$$

Then the post-exchange decisions are still in the reference interval  $[-1, 1]$  if we assume  $0 \leq P(\cdot, \cdot; \cdot, \cdot) \leq 1$  and  $\alpha_S + 2\alpha_P \leq 2$ . We can state the following result which identifies the condition on the noise term in order to ensure that the post-interaction opinions do not leave the reference interval.

**Proposition 1.** *We assume  $0 < P(x, x_*; w, w_*) \leq 1$ ,  $S(\cdot) = S \leq 1$  and*

$$|\eta| < (1 - \alpha_P + \alpha_S)d, \quad |\eta_*| < (1 - \alpha_P + \alpha_S)d, \quad \alpha_P \leq 1,$$

where

$$d = \min_{(x, w) \in X \times I} \left\{ \frac{(1 - w)}{D(x, w)}, D(x, w) \neq 0 \right\}.$$

Then the binary interaction rule (2.7) preserves the bounds being  $w', w'_* \in [-1, 1]$ .

*Proof.* The proof is analogous to those in [1, 2, 11], therefore we omit the details.  $\square$

The introduction of a drift term towards the correct decision in equation (2.7) follows from the set-up of the two-alternative forced choice tasks (TAFC) addressed in [5, 30]. The (TAFC) task is a classic behavioral experiment: at each trial agents are called to recognize a noisy visual stimulus choosing between two alternatives. The individual decision making involves in its mathematical description a drift term representing the fact that increasing experience/competence drives the agent towards the correct decision.

The compromise function  $0 \leq P(\cdot, \cdot; \cdot, \cdot) \leq 1$  depends on both the particles' opinion and competence. As an example, a possible structure for the interaction function is given by the following

$$P(w, w_*; x, x_*) = Q(w, w_*)R(x, x_*), \quad (2.9)$$

where  $0 \leq Q(\cdot, \cdot) \leq 1$  represents the positive compromise propensity and  $0 \leq R(\cdot, \cdot) \leq 1$  is a function taking into account the competence of two interacting agents.

Empirical experiments have been done in order to measure the impact of competence on collective decision-making tasks [4, 22, 24]. Here the participants are organized in dyads which are called to make a decision about a visual stimulus. Among others two opposite models taking into account the influence of the competence, or *sensitivity* in the psychology literature, might be considered. The first proposes that nothing except the opinions is communicated between individuals, and in case of disagreement we randomly select an opinion which coincides with the decision of the studied dyad. We will refer to this model as coin-flip model (CF). The second model takes strongly into account the competence of individuals: the decisions are communicated and in case of disagreement the opinion of the most competent prevails and then coincides with the decision of the dyad. This second model is called maximum competence model (MC).

The aforementioned models can be described in our mathematical setting by considering a compromise propensity  $Q(w, w_*) = 1$ . The CF model is a model in which at each interaction is

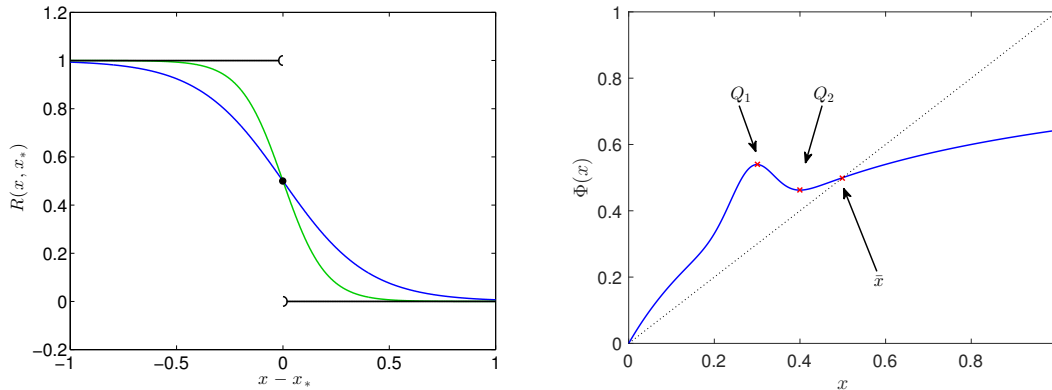


Figure 1: Left: Heaviside-type competence-based interaction function  $R_{MC}(x, x_*)$  and its continuous version defined in (2.11) for  $c = 5$  and  $c = 10$ . Right: equality bias function  $\Phi(x)$  of (2.14), the choice of coefficients  $a_1, a_2, a_3, a_4$  has been done such that  $x_{Q_1} = 0.3, x_{Q_2} = 0.4$  and  $\bar{x} = 0.5$ .

associated a Bernoulli random variable with typical parameter  $p = 1/2$ . In other words, in the CF case, the competence does not play any role in the decision process. We will consider an averaged version of this model (aCF), which corresponds to  $R_{aCF}(\cdot, \cdot) = 1/2$ . For the MC model, in our setting corresponds to the Heaviside-type function

$$R_{MC}(x, x_*) = \begin{cases} 1 & x < x_* \\ 1/2 & x = x_* \\ 0 & x > x_*. \end{cases} \quad (2.10)$$

In terms of the competence gap  $x - x_*$  we can approximate the discontinuous function (2.10) of the MC model with a smoothed continuous version (cMC)

$$R_{cMC}(x, x_*) = \frac{1}{1 + e^{c(x-x_*)}}, \quad (2.11)$$

with  $c \gg 1$ . We depict in Figure 1, left plot, the behavior of the function  $R(\cdot, \cdot)$  for different choices of the constant  $c > 0$ . We can observe how in the half-plane  $x > x_*$  the most competent agent is scarcely influenced by the less skilled one, while in the half-plane  $x < x_*$  the situation is inverted and we see how an agent with less competence is influenced by the more competent one.

### 2.3 Dynamics of decisions under equality bias

On the basis of the decision process built in section 2.2 we modify the microscopic model (2.7)–(2.9) in order to mathematically describe the phenomenon called *equality bias* in collective decision-making communities. Our set-up is essentially inspired by some recent works of the experimental psychology literature [4, 24] and the references therein. Their findings are consistent with the well-known cognitive bias called Dunning-Kruger effect regarding a misjudgment of personal competence of unskilled people, which overestimate their own ability. At the same time the most skilled individuals tends to underestimate their competence, implicitly believing that their knowledge is accessible to everyone [22].

Let us consider an agents with competence  $x \in X$ . We introduce the equality bias function

$$\begin{aligned}\Phi : X &\mapsto X \\ x &\rightarrow \Phi(x)\end{aligned}\tag{2.12}$$

which measures the agent's perceived level of competence. The choice of the equality bias function  $\Phi(\cdot)$  is related to the subjective self-confidence in making decisions and experimental results show it is strictly dependent on the degree of competence of the interacting agents.

Following the mathematical setting introduced in section 2.2 we modify the function  $P(\cdot, \cdot; \cdot, \cdot)$  as

$$P(\Phi(x), \Phi(x_*); w, w_*).\tag{2.13}$$

In particular in the case (2.9) the competence based interaction function modifies in  $R(\Phi(x), \Phi(x_*))$  whose shape depends now on the perceived competence. Therefore the equality bias emerges as a result of the binary interactions of a multi-agent system with lack of metacognition. In agreement with the experimental measurements of [22], and in order to exemplify the action of the equality bias function, we depict in Figure 1 the function

$$\Phi(x) = x \left( \frac{a_1}{x + a_2} + a_3 \exp \{-h(x - a_4)^2\} \right),\tag{2.14}$$

with coefficients  $a_1, a_2, a_3, a_4 > 0$ . In the presented equality bias function we fixed the value  $x_d \in X$  measuring the average competence of a population, the agents with competence  $x > x_d$  may be identified with the most competent. Moreover, as experimentally suggested, it exists a local maximum  $Q_1$  in  $x = x_{Q_1}$  for the overestimation of the competence of unskilled agents, together with the local minimum  $Q_2$  in  $x = x_{Q_2}$ . The asymptotic perceived competence depends on  $a_1$ , whose speed is weighted by  $a_2$ . Therefore we choose the coefficients  $a_2, a_3, a_4$  in such a way that  $\Phi(x_d) = x_d$  and for  $x = x_{Q_1}, x = x_{Q_2}$  we find a local maximum and minimum respectively. The coefficient  $h > 0$  is a scaling parameter. In the rest of the paper we will refer to the dynamics of decisions under the action of the equality bias with (EB).

### 3 A Boltzmann model for opinion and competence

In this section we derive a kinetic model for opinion and competence reflecting the behavior introduced in the binary interaction model for opinion and competence in (2.2)–(2.7).

Let  $f = f(x, w, t)$  be the density of individuals with competence  $x \in X \subset \mathbb{R}^+$  and opinion  $w \in I = [-1, 1]$  at time  $t \geq 0$ . We derive the kinetic description for the evolution of the density function  $f = f(x, w, t)$  through classic methods of kinetic theory [29]. Let  $g(x, t)$  be the marginal density of the competence variable  $x \in X$

$$g(x, t) = \int_I f(x, w, t) dw,\tag{3.1}$$

and  $\int_X g(x, t) = 1$  for each  $t \geq 0$ .

The evolution in time of the introduced density function is then given by the integro-differential equation of the Boltzmann-type

$$\frac{\partial}{\partial t} f(x, w, t) = Q(f, f)(x, w, t),\tag{3.2}$$

where  $f(x, v, 0) = f_0(x, w)$  and  $Q(\cdot, \cdot)$  is defined as follows

$$Q(f, f)(x, w, t) = \int_{\mathbb{B}^2} \int_{X \times I} \left( {}'B \frac{1}{J} f({}'x, {}'w, t) f({}'x_*, {}'w_*, t) - B f(x, w, t) f(x_*, w_*, t) \right) dw_* dx_* d\eta d\eta_*,\tag{3.3}$$

indicating with  $(w, w_*)$  the pre-interaction opinions given by  $(w, w_*)$  after the interaction and  $(x, x_*)$  the pre-interaction competences. The term  $J = J(x, w; x_*, w_*)$  denotes as usual the Jacobian of the transformations  $(w, w_*) \rightarrow (w', w'_*)$ ,  $(x, x_*) \rightarrow (x', x'_*)$  via (2.2) and (2.7). The kernels  $B, B'$  characterize the binary interaction and in the following will be considered of the form

$$B_{(w, w_*) \rightarrow (w', w'_*)}^{(x, x_*) \rightarrow (x', x'_*)} = \beta \Theta(\eta) \Theta(\eta_*) C(z) \chi(|w'| \leq 1) \chi(|w'_*| \leq 1) \chi(x' \in X) \chi(x'_* \in X), \quad (3.4)$$

where  $\beta > 0$  is a scaling constant.

The presence of the Jacobian in the definition of the binary operator (3.3) may be avoided by considering its weak formulation. Let us consider a test function  $\psi(x, w)$ , we get

$$\begin{aligned} & \int_{X \times I} Q(f, f)(x, w, t) \psi(x, w) dw dx \\ &= \beta \left\langle \int_{\mathbb{R}^+} \int_{X^2 \times I^2} (\psi(x', w') - \psi(x, w)) f(x, w, t) f(x_*, w_*, t) C(z) dw dw_* dx dx_* dz \right\rangle \\ &= \frac{\beta}{2} \left\langle \int_{\mathbb{R}^+} \int_{X^2 \times I^2} (\psi(x', w') + \psi(x'_*, w'_*) - \psi(x, w) - \psi(x_*, w_*)) \right. \\ & \quad \left. f(x, w, t) f(x_*, w_*, t) C(z) dw dw_* dx dx_* dz \right\rangle \end{aligned} \quad (3.5)$$

where the brackets  $\langle \cdot \rangle$  denotes the expectation with respect to the random variables  $\eta, \eta_*$ . The weak formulation of the initial value problem (3.2) for the initial density  $f_0(x, w)$  is given for each  $t \geq 0$  by

$$\begin{aligned} & \frac{\partial}{\partial t} \int_{X \times I} \psi(x, w) f(x, w, t) dw dx \\ &= \frac{\beta}{2} \left\langle \int_{\mathbb{R}^+} \int_{X^2 \times I^2} (\psi(x', w') + \psi(x'_*, w'_*) - \psi(x, w) - \psi(x_*, w_*)) \right. \\ & \quad \left. f(x, w, t) f(x_*, w_*, t) C(z) dw dw_* dx dx_* dz \right\rangle. \end{aligned} \quad (3.6)$$

From the weak formulation in (3.6) we can derive the evolution of the macroscopic quantities like the moments for the opinion which may be obtained choosing as a test function  $\psi(x, w) = \psi(w) = 1, w, w^2$ .

### 3.1 Collective decision and variance

It is straightforward to observe that setting  $\psi = 1$  we obtain the conservation of the total number of agents. The mean opinion of the overall agents is defined as

$$U(t) = \int_{X \times I} w f(x, w, t) dw dx, \quad (3.7)$$



which represents the *collective decision* of the system at time  $t$ , see [9, 16, 17]. The evolution of the collective decision is derived as marginal quantity from equation (3.6) for  $\psi(x, w) = w$

$$\begin{aligned} \frac{\partial}{\partial t} \int_{X \times I} w f(x, w, t) dw dx &= \alpha_S \beta \int_{X \times I} S(x) (w_d - w) f(x, w, t) dw dx \\ &+ \frac{\alpha_P \beta}{2} \int_{X^2 \times I^2} (P(x, x_*; w, w_*) - P(x_*, x; w_*, w)) (w - w_*) \\ &f(x, w, t) f(x_*, w_*, t) dw dw_* dx dx_*. \end{aligned} \quad (3.8)$$

If the interaction function  $P(\cdot, \cdot; \cdot, \cdot)$  is a symmetric function and  $S(x) = s \in [0, 1]$  equation (3.8) reduces to

$$\frac{d}{dt} U(t) = \alpha_S \beta s (w_d - U(t)), \quad (3.9)$$

whose solution at each  $t \geq 0$  is  $U(t) = w_d + (U_0 - w_d) \exp\{-\alpha_S \beta s t\}$ , with  $U_0 = U(0)$  the initial collective decision. Therefore, in the limit  $t \rightarrow +\infty$ , the asymptotic collective decision converges exponentially toward  $w_d \in \{-1, 1\}$ , i.e.  $U_\infty = w_d$ .

In the case of the aCF model  $P \equiv 1/2$  and  $S(\cdot) \equiv 0$ , the mean opinion of the overall system is then conserved, which implies  $U(t) = U_0$  for all  $t \geq 0$ . Further if we consider a MC model, with interactions between agents based only on the competence variable and for  $S \equiv 0$ , from (3.8) we have

$$\begin{aligned} \frac{d}{dt} \int_{X \times I} w f(x, w, t) dx dw &= \alpha_P \beta \int_{X^2 \times I^2} R_{MC}(x, x_*) (w_* - w) f(x, w, t) f(x_*, w_*, t) dw dw_* dx dx_* \\ &- \frac{\alpha_P \beta}{2} \int_{X^2 \times I^2} (w_* - w) f(x, w, t) f(x_*, w_*, t) dw dw_* dx dx_*, \end{aligned} \quad (3.10)$$

being  $R_{MC}(x, x_*) - R_{MC}(x_*, x) = 2R_{MC}(x, x_*) - 1$  with  $R_{MC} \equiv 0$  in the half space  $x > x_*$ , which leads to

$$\begin{aligned} \frac{d}{dt} \int_X u(x, t) dx &= \alpha_P \beta \int_X \left[ \frac{U^+(x, t) - U^-(x, t)}{2} \right] \rho(x, t) dx \\ &+ \alpha_P \beta \int_X \left[ \frac{\rho^-(x, t) - \rho^+(x, t)}{2} \right] u(x, t) dx \end{aligned} \quad (3.11)$$

where  $u(x, t) = \int_I w f(x, w, t) dw$  is the mean opinion relative to the competence level  $x \in X$  and

$$U^+(x, t) = \int_{x < x_*} \int_I w_* f(x_*, w_*, t) dw_* dx_*; \quad \rho^+(x, t) = \int_{x < x_*} \int_I f(x_*, w_*, t) dw_* dx_*. \quad (3.12)$$

In particular  $U^+(x, t)$  and  $\rho^+(x, t)$  indicate the average opinion and the numerical density of agents with competence greater than  $x \in X$ . We define also  $U^-(x, t) = U(t) - U^+(x, t)$  and  $\rho^-(x, t) = 1 - \rho^+(x, t)$  for each  $x \in X$  and  $t \geq 0$ . Therefore the variations of the mean opinion of agents with fixed competence  $u(x, t)$  follow the choice of the most competent agents.

The second order moment for the opinion of the overall system is defined as

$$E(t) = \int_{X \times I} w^2 f(x, w, t) dw dx \quad (3.13)$$

and its evolution may be obtained from (3.6) with  $\psi(x, w) = w^2$

$$\begin{aligned}
\frac{d}{dt}E(t) &= \alpha_S^2 \beta \int_{X \times I} S^2(x) (w - w_d)^2 f(x, w, t) dw dx \\
&+ \frac{\alpha_P^2 \beta}{2} \int_{X^2 \times I^2} [P^2(x, x_*; w, w_*) + P^2(x_*, x; w_*, w)] (w - w_*)^2 f(x, w, t) f(x_*, w_*, t) dw dw_* dx dx_* \\
&- \alpha_P \beta \int_{X^2 \times I^2} (w - w_*) [w P(x, x_*; w, w_*) - w_* P(x_*, x; w_*, w)] f(x, w, t) f(x_*, w_*, t) dw dw_* dx dx_* \\
&- 2\alpha_S \beta \int_{X \times I} S(x) w (w - w_d) f(x, w, t) dw dx \\
&+ \alpha_S \alpha_P \beta \int_{X^2 \times I^2} (w - w_*) [S(x) P(x, x_*; w, w_*) (w - w_d) - S(x_*) P(x_*, x; w_*, w) (w_* - w_d)] \\
&\quad f(x, w, t) f(x_*, w_*, t) dw dw_* dx dx_* + \alpha_P \beta \sigma^2 \int_{X \times I} D^2(x, w) f(x, w, t) dw dx.
\end{aligned} \tag{3.14}$$

In the simplified situation  $P(x, x_*; w, w_*) = p \in [0, 1]$ ,  $S(x) = s \in [0, 1]$ ,  $\alpha_S = \alpha_P = \alpha$  and in absence of diffusion, the equation (3.14) assumes the following form

$$\begin{aligned}
\frac{d}{dt}E(t) &= \alpha^2 \beta s^2 \int_{X \times I} (w - w_d)^2 f(x, w, t) dw dx - 2\alpha \beta s \int_{X \times I} (w^2 - w w_d) f(x, w, t) dw dx \\
&+ \alpha \beta p (\alpha p - 1 + \alpha s) \int_{X^2 \times I^2} (w - w_*)^2 f(x, w, t) f(x_*, w_*, t) dw dw_* dx dx_*
\end{aligned} \tag{3.15}$$

Since under the same conditions  $U(t)$  converges for large time toward the correct choice  $w_d \in \{-1, 1\}$ , we obtain that  $E(t)$  converges exponentially to  $w_d^2$  if

$$\alpha \leq \min \left\{ 1, \frac{2(s+p)}{(s+p)^2 + p^2} \right\}. \tag{3.16}$$

Therefore the quantity

$$\int_{X \times I} (w - w_d)^2 f(x, w, t) dw dx, \tag{3.17}$$

under the above assumptions, converges toward zero for large times and the steady state solution is the Dirac delta  $\delta(w - w_d)$  centered in the right decision.

## 4 Fokker-Planck approximation

In order to obtain analytic results on the large-time behavior from Boltzmann-type models a classical mathematical tool is given by the derivation of approximated Fokker-Planck models through scaling techniques [1, 2, 3, 7, 15, 28, 29]. In what follows we apply this approach, also known as quasi-invariant limit [32, 29], to the model derived in the latter section.

We introduce a scaling parameter  $\epsilon > 0$  and the following scaled quantities

$$\lambda = \epsilon \lambda, \quad \lambda_B = \epsilon \lambda_B, \quad \lambda_C = \epsilon \lambda_C, \quad \sigma_\kappa = \sqrt{\epsilon} \sigma_\kappa, \tag{4.1}$$

where we assumed  $\lambda_B(x) = \lambda_B > 0$ ,  $\lambda_C(x) = \lambda_C > 0$  and  $\lambda(x) = \lambda > 0$ . This corresponds to the situation where each interaction produces a small variation of the competence. The same strategy

may be applied to the binary opinion model rescaling the interaction frequency  $\beta$ , the interaction propensity  $\alpha$  and the diffusion variance  $\sigma^2$  as follows

$$\alpha_S = \epsilon \alpha_S, \quad \alpha_P = \epsilon \alpha_P, \quad \beta = \frac{1}{\epsilon}, \quad \sigma = \sqrt{\epsilon} \sigma. \quad (4.2)$$

The scaled equation (3.6) reads

$$\begin{aligned} \frac{\partial}{\partial t} \int_{X \times I} \psi(x, w) f(x, w, t) dw dx = \\ \frac{1}{\epsilon} \left\langle \int_{\mathbb{R}^+} \int_{X^2 \times I^2} (\psi(x', w') - \psi(x, w)) f(x, w, t) f(x_*, w_*, t) C(z) dw dw_* dx dx_* dz \right\rangle. \end{aligned} \quad (4.3)$$

Under the assumptions on the random variables involved in the binary exchanges  $\kappa, \eta, \eta_*$  we define the following mean quantities

$$\begin{aligned} \langle x' - x \rangle &= -\lambda x + \lambda_C x_* + \lambda_B z = G(x, x_*, z), \\ \langle w' - w \rangle &= -\alpha_S S(x)(w - w_d) - \alpha_P P(x, x_*; w, w_*)(w - w_*) \\ &= H_S(x, w) + H_P(x, x_*; w, w_*), \\ \langle (x' - x)^2 \rangle &= G^2(x, x_*, z) + \sigma_\kappa^2 x^2, \\ \langle (w' - w)^2 \rangle &= H^2(x, x_*; w, w_*) + \sigma^2 D^2(x, w), \\ \langle (x' - x)(w' - w) \rangle &= G(x, x_*, z) H(x, x_*; w, w_*). \end{aligned} \quad (4.4)$$

Then, we have

$$\begin{aligned} \langle \psi(x', w') - \psi(x, w) \rangle &= G(x, x_*, z) \frac{\partial \psi}{\partial x}(x, w) + (H_S(x, w) + H_P(x, x_*; w, w_*)) \frac{\partial \psi}{\partial w}(x, w) \\ &\quad \frac{1}{2} \left[ (G^2(x, x_*, z) + \sigma_\kappa^2 x^2) \frac{\partial^2 \psi}{\partial x^2}(x, w) + ((H_S(x, w) + H_P(x, x_*; w, w_*))^2 + \sigma^2 D^2(x, w)) \frac{\partial^2 \psi}{\partial w^2}(x, w) \right. \\ &\quad \left. G(x, x_*, z)(H_S(x, w) + H_P(x, x_*; w, w_*)) \frac{\partial^2 \psi}{\partial x \partial w}(x, w) \right] + R(x, x_*; w, w_*), \end{aligned} \quad (4.5)$$

where  $R(x, x_*; w, w_*)$  denotes the higher order terms of the Taylor expansion. From the quasi-invariant scalings introduced in (4.1)-(4.2) it follows that

$$G(x, x_*, z) = \epsilon \tilde{G}(x, x_*, z), \quad H_S(x, w) = \epsilon \tilde{H}_S(x, w), \quad H_P(x, x_*; w, w_*) = \epsilon \tilde{H}_P(x, x_*; w, w_*). \quad (4.6)$$

Then equation (4.3) takes the form

$$\begin{aligned} \frac{\partial}{\partial t} \int_{X \times I} \psi(x, w) f(x, w, t) dw dx = \\ \int_{\mathbb{R}^+} \int_{X^2 \times I^2} \left[ \tilde{G}(x, x_*, z) \frac{\partial \psi}{\partial x}(x, w) + \tilde{H}_S(x, w) \frac{\partial \psi}{\partial w}(x, w) + \tilde{H}_P(x, x_*; w, w_*) \frac{\partial \psi}{\partial w}(x, w) \right. \\ \left. + \frac{\sigma_\kappa^2}{2} x^2 \frac{\partial^2 \psi}{\partial x^2}(x, w) + \frac{\sigma^2}{2} D^2(x, w) \frac{\partial^2 \psi}{\partial w^2}(x, w) \right] f(x, w, t) f(x_*, w_*, t) C(z) dw dw_* dx dx_* dz \\ + \tilde{R}(\epsilon) + O(\epsilon), \end{aligned} \quad (4.7)$$

with

$$\begin{aligned}
\tilde{R}(\epsilon) = & \frac{1}{2\epsilon} \int_{\mathbb{R}^+} \int_{X^2 \times I^2} \left[ \epsilon^2 \left( \tilde{G}^2(x, x_*, z) \frac{\partial^2 \psi}{\partial x^2}(x, w) \right. \right. \\
& + (\tilde{H}_S(x, w) + \tilde{H}_P(x, x_*; w, w_*))^2 \frac{\partial^2 \psi}{\partial w^2}(x, w) \\
& + \tilde{G}(x, x_*, z) (\tilde{H}_S(x, w) + \tilde{H}_P(x, x_*; w, w_*)) \frac{\partial^2 \psi}{\partial x \partial w}(x, w) \Big) \\
& \left. + R(x, x_*; w, w_*) \right] f(x, w, t) f(x_*, w_*, t) C(z) dw dw_* dx dx_* dz.
\end{aligned} \tag{4.8}$$

By similar arguments to [28, 32] it can be shown that the term  $R(\epsilon)$  defined in (4.8) decays to zero in the limit  $\epsilon \rightarrow 0$ . Finally for  $\epsilon \rightarrow 0$  we obtain

$$\begin{aligned}
\frac{\partial}{\partial t} \int_{X \times I} \psi(x, w) f(x, w, t) dw dx = & \int_{\mathbb{R}^+} \int_{X^2 \times I^2} \left[ \tilde{G}(x, x_*, z) \frac{\partial \psi}{\partial x}(x, w) \right. \\
& + \tilde{H}_S(x, w) \frac{\partial \psi}{\partial w}(x, w) + \tilde{H}_P(x, x_*; w, w_*) \frac{\partial \psi}{\partial w}(x, w) \Big] f(x, w, t) f(x_*, w_*, t) C(z) dw dw_* dx dx_* dz \\
& + \frac{\sigma_\kappa^2}{2} \int_{X^2 \times I^2} x^2 \frac{\partial \psi}{\partial x^2}(x, w) f(x, w, t) dw dx \\
& + \frac{\sigma^2}{2} \int_{X^2 \times I^2} D^2(x, w) \frac{\partial^2 \psi}{\partial w^2}(x, w) f(x, w, t) dw dx.
\end{aligned} \tag{4.9}$$

Integrating back by parts equation (4.9) we have the following nonlinear Fokker-Planck equation

$$\begin{aligned}
\frac{\partial}{\partial t} f(x, w, t) = & \frac{\partial}{\partial x} \mathcal{G}[f](x, t) f(x, w, t) + \frac{\partial}{\partial w} \mathcal{H}[f] f(x, w, t) + \frac{\partial}{\partial w} \mathcal{K}[f] f(x, w, t) \\
& + \frac{\sigma_\kappa^2}{2} \frac{\partial^2}{\partial x^2} (x^2 f(x, w, t)) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial w^2} (D^2(x, w) f(x, w, t)),
\end{aligned} \tag{4.10}$$

where

$$\begin{aligned}
\mathcal{G}[f](x, t) = & \int_{\mathbb{R}^+} \int_{X \times I} (\lambda x - \lambda_C x_* - \lambda_B z) f(x_*, w_*, t) C(z) dw_* dx_* dz \\
= & \lambda x - \lambda_C m_x - \lambda_B m_B,
\end{aligned} \tag{4.11}$$

being  $m_x(t) = \int_{X \times I} x f(x, w, t) dw dx$ , and where the functionals  $\mathcal{H}[f], \mathcal{K}[f]$  are defined as follows

$$\begin{aligned}
\mathcal{H}[f](x, w, t) = & \int_{X \times I} \alpha_P P(x, x_*; w, w_*) (w - w_*) f(x_*, w_*, t) dw_* dx_* \\
\mathcal{K}[f](w_d) = & \int_{X \times I} \alpha_S S(x) (w - w_d) f(x, w, t) dw dx.
\end{aligned} \tag{4.12}$$

#### 4.1 Stationary states for the marginal density

We introduce the marginal density function of the competence

$$g(x, t) = \int_I f(x, w, t) dw. \tag{4.13}$$

By direct integration of the Fokker-Planck equation (4.10) with respect to the variable  $w$  and considering a local diffusion function  $D(x, w)$  decaying at the boundaries of the reference interval for the opinions, and taking into account the convergence of the mean competence showed in Section 2.1, we obtain

$$\frac{\partial}{\partial t} g(x, t) = \frac{\partial}{\partial x} \left( \lambda x - \frac{\lambda \lambda_B}{\lambda - \lambda_C} m_B \right) g(x, t) + \int_I \mathcal{H}[f] f(x, w, t) dw + \frac{\sigma_\kappa^2}{2} \frac{\partial^2}{\partial x^2} (x^2 g(x, t)). \quad (4.14)$$

The above equation simplifies by imposing the Dirichlet boundary conditions  $f(x, -1, t) = f(x, 1, t) = 0$  for each  $x \in X, t \geq 0$ . Therefore it is possible to give the analytic formulation of the stationary solution of (4.14). The solution of

$$\frac{\partial}{\partial x} \left( -\lambda x + \frac{\lambda \lambda_B}{\lambda - \lambda_C} m_B \right) g^\infty(x) = \frac{\sigma_\kappa^2}{2} \frac{\partial^2}{\partial x^2} (x^2 g^\infty(x)) \quad (4.15)$$

is given by

$$g^\infty(x) = \frac{c_{\lambda, \lambda_B, \lambda_C, \sigma_\kappa^2}}{x^{2+2\lambda/\sigma_\kappa^2}} \exp \left\{ -\frac{2}{\sigma_\kappa^2 x} \cdot \frac{\lambda \lambda_B m_B}{\lambda - \lambda_C} \right\}, \quad (4.16)$$

where  $c_{\lambda, \lambda_B, \lambda_C, \sigma_\kappa^2}$  is a constant chosen such that the total mass of  $g^\infty$  is equal to one.

Unlike the usual method for determining the stationary density developed in [7, 28], here the asymptotic competence has been derived from the complete Fokker-Planck equation (4.10) after the integration of the opinion variable, and under specific boundary conditions.

## 5 Numerics

In this section we propose several numerical tests for the Boltzmann-type model introduced in the previous paragraphs which show the emergence of suboptimal collective decisions under the hypothesis of equality bias compared to aCF and cMC models. All the results presented have been obtained through direct simulation Monte Carlo methods for the Boltzmann equation (see [27, 29]).

The Fokker-Planck regime is obtained via the quasi-invariant scaling (4.1)–(4.2) with  $\epsilon = 0.01$ . The local diffusion function has been considered in the case

$$D(x, w) = 1 - w^2 \quad (5.1)$$

which multiplies in the binary collision model (2.7) for the opinion, a uniform random variable with scaled variance  $\sigma^2$ . The choice (5.1) implies that the diffusion does not act on the agents with more extremal opinions. In the competence dynamics we considered  $\lambda(x) = \lambda$ ,  $\lambda_C(x) = \lambda_C$  and  $\lambda_B(x) = \lambda_B$  and a uniform random variable  $\kappa$  with finite scaled variance  $\sigma_\kappa^2$ . The random variable  $z$  in here uniform with mean  $1/2$ .

### 5.1 Test 1: collective decision under equality bias

In this test we compute the collective decisions emerging from the Boltzmann-type model (3.2) with interaction function (2.9) and  $Q(w, w_*) = 1$  and  $R(x, x_*)$  defined in the reference cases aCF-cMC and EB. The binary interaction terms defined in (2.2)–(2.7) have been considered for a choice of constants coherent with the bounds described in Section 2. In this test we considered a vanishing drift term  $S(x) = 0$ , for each  $x \in X$ .

We compare the emerging collective decisions in the aCF and cMC cases considering the action of the equality bias function described in equation (2.14). In this test we take into account the compromise behavior of two strongly polarized populations with equal sizes (case A) and with

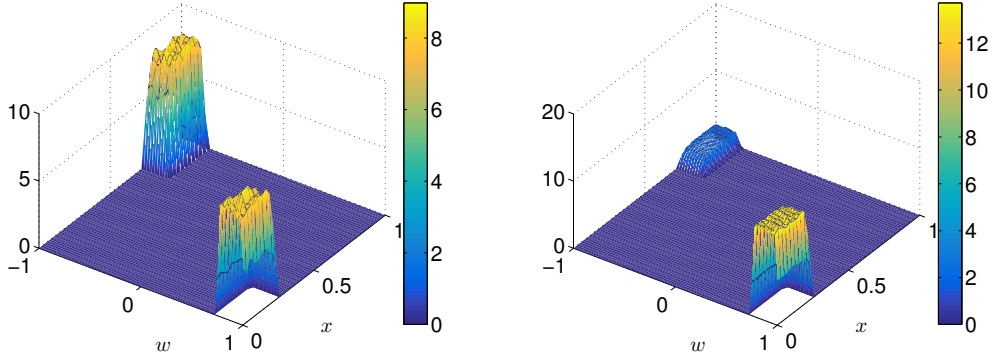


Figure 2: Left: **Test 1A**, initial configuration of the multi-agent system with an equal size of competent and incompetent agents. Right: **Test 1B**, initial configuration of the multi-agent system with a different size of competent and incompetent agents, in particular the number of competent agents is five times smaller than the size of the incompetent agents.

different sizes (case B). In both cases the group of competent agents is characterized at  $t = 0$  by opinions in the interval  $w \in [-1, -0.75]$ , whereas the second population, composed by the less competent agents, expresses opinions  $w \in [0.75, 1]$ . In the case of populations with different size we considered an initial distribution such that the number of competent agents is five times larger than the number of the incompetent ones. In Figure 2 we exemplified the initial configurations for the two tests: Test 1A and Test 1B.

In Figure 4 we show the kinetic and particle solutions of the Test 1A for a system of interacting agents in the aCF-cMC and EB1-EB2 models, where the case of equality bias has been considered for two examples of equality bias functions  $\Phi_1(x)$  and  $\Phi_2(x)$  of the type introduced in Section 2.3, see equation (2.14), and whose behaviors are reported in Figure 3. Further in Figure 3 we compare the stationary distribution for the competence variable emerging from the Monte Carlo method and its analytic formulation obtained in Section 4.1. In particular in Figure 4 we can observe how the emerging collective decisions of the EB1 and EB2 cases are significantly shifted from the decision emerging in a model cMC.

In Figure 5 we show the stationary distributions of the opinion variable in the four models of interactions both the Test 1A (left) and Test 1B (right). We observe how the EB1-EB2 models in general defines a decision which is suboptimal with respect to the competence based model cMC. Further in presence of a strong overestimation of the opinions of the less skilled agents, like in the EB2 model, we see how the emerging decision may perform worse than a aCF mode, i.e. a model which does not take into account the competence at all. In the case of the Test 1B we see how for asymmetric populations the emerging decision, also in the cMC model, may be deviated toward the positions of the less competent agents. The same behavior is highlighted in Figure 6 where we find the evolution of the mean opinion of the multi-agent systems which lead to the formation of the collective decisions.

## 5.2 Test 2: competence driven optimal decision

We consider here the action of the term driving the system to the correct decision introduced in the opinion dynamics (2.7). There we included a competence-dependent force with a rate given by  $S(\cdot)$  representing an increasing evidence in supporting the a-priori right choice  $w_d \in \{-1, 1\}$ . We

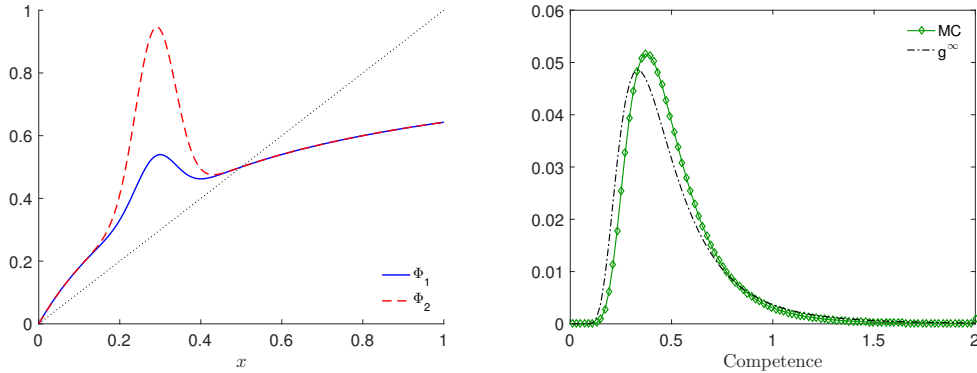


Figure 3: **Test 1.** Left: two examples of equality bias functions  $\Phi(x)$  if the form introduced in (2.14) and passing through  $\bar{x} = 0.5$  with  $a_1 = 0.9$  and scaling parameter  $h = 10^2$ . In the rest of the paper we will refer to them as  $\Phi_1(x)$  and  $\Phi_2(x)$ . Right: convergence of the Monte Carlo method toward the analytic steady state for  $\epsilon = 0.01$ .

consider here

$$S(x) = \begin{cases} \min\{1, x\} & x \geq x_d \\ 0 & x \leq x_d \end{cases} \quad (5.2)$$

and  $\lambda_C = \lambda_C \chi(x \geq x_d)$ . In the proposed set-up we intended to mimic the fact that extremely low skilled people ( $x \leq x_d$ ), in addition to make wrong choices, have not the ability to realize the inaccuracy of their decision, a phenomenon which follows from the Dunning-Kruger effect. In Figure 7 we report the evolution of the multi-agent system with the same initial configuration as in Test 1B. The forcing term is characterized by (5.2) with  $x_d = 0.3$ , which drives the opinions of sufficiently competent agents toward  $w_d = -1$ . Here the equality bias functions  $\Phi_1(x)$  and  $\Phi_2(x)$ , introduced in 2, influences the speed of convergence of the opinions toward  $w_d$ . We compare in Figure 8 the convergences toward  $w_d$  for the models aCF-cMC and EB1-EB2.

### 5.3 Test 3: bounded confidence case

In the last test case we consider the case of bounded confidence based interactions [1, 3, 6]. Often we notice how more well-educated, more capable and more competent people are also those best disposed to dialogue. Then competence is generally associated to the predisposition to listen other people. The higher this quality, greater is the ability to value other opinions. Vice versa, a person unwilling to listen and dialogue is usually marked by a lower level of the described trait. Therefore, it is natural to consider a bounded confidence model where the threshold on the exchanges of opinions depends on the degree of competence.

In particular, we consider a compromise function  $P(x, x_*; w, w_*)$  with the following form

$$P(x, x_*; w, w_*) = \chi(|w - w_*| \leq \gamma \Delta(x, x_*)), \quad (5.3)$$

where  $\Delta(x, x_*)$  is a competence-dependent function that ranges in the closed interval  $[0, 1]$  taking into account the maximum distance under which the interactions are allowed, and  $\gamma > 0$  is a constant value. A possible choice of for the function  $\Delta(\cdot, \cdot)$  is

$$\Delta(x, x_*) = R_{cMC}(x, x_*), \quad (5.4)$$

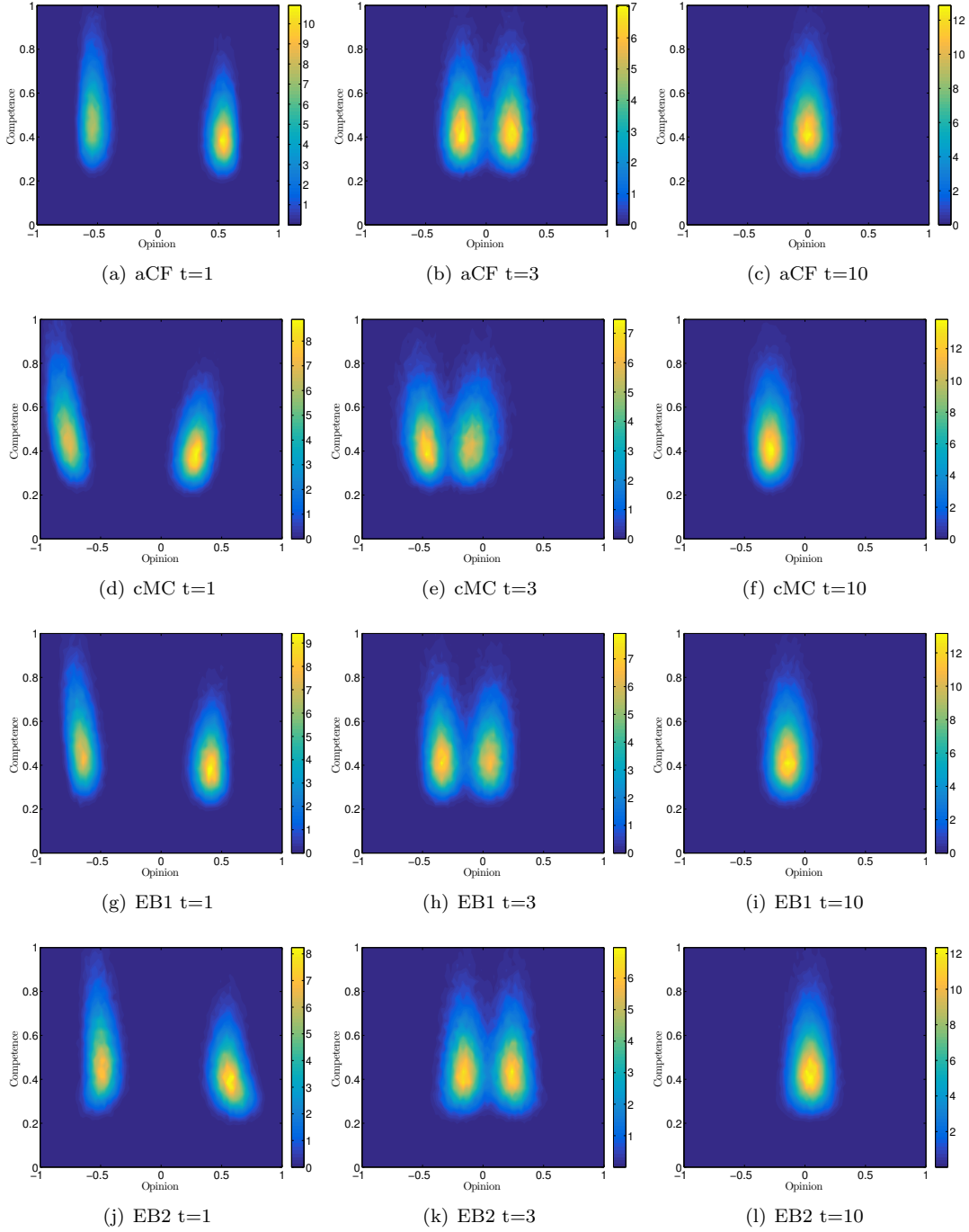


Figure 4: **Test 1A**: kinetic solution at different time steps in the aCF-cMC and EB1-EB2 model cases, respectively from the first row. The choice of constants in this test is:  $c = 10$ ,  $\lambda_B = \lambda_C = 10^{-2}$ ,  $\sigma_\kappa^2 = 10^{-2}$  and  $\lambda = \lambda_B + \lambda_C$ . The mean competence of the multi-agent system is  $\bar{x} = 0.5$ . We see how the equality bias, through the equality bias functions  $\Phi_1(x)$  and  $\Phi_2(x)$  presented in Figure 3, influences the opinion dynamics, driving the collective decision toward suboptimal states with respect to the cMC model.



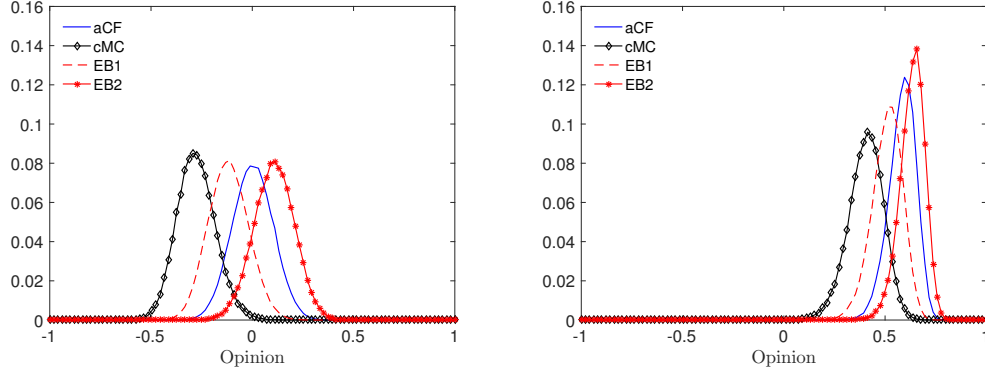


Figure 5: Asymptotic distributions of the opinion variable for the models aCF-cMC-EB1-EB2 Left: **Test 1A**, case of initial configuration with equal size of competent and incompetent agents. Right: **Test1B**, case of asymmetric populations as in Figure 2. The choice of constants is  $c = 10$ ,  $\lambda_B = \lambda_C = 10^{-2}$ ,  $\sigma_\kappa^2 = 10^{-2}$  and  $\lambda = \lambda_B + \lambda_C$ , the equality bias functions are sketched in Figure 3.

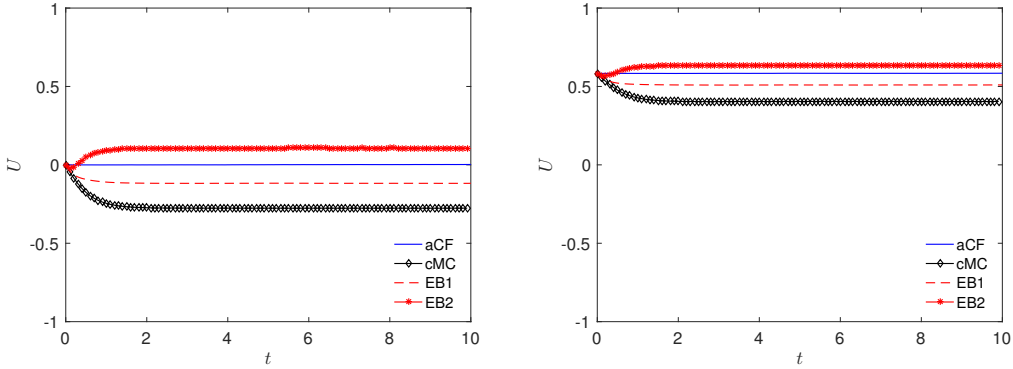


Figure 6: Collective decisions of the multi-agent system for the reference models aCF-cMC-EB1-EB2 with the choice of parameters  $c = 10$ ,  $\lambda_B = \lambda_C = 10^{-2}$ ,  $\sigma_\kappa^2 = 10^{-2}$  and  $\lambda = \lambda_B + \lambda_C$ ; the variable  $z$  is . Left: **Test 1A**. Right: **Test1B**.

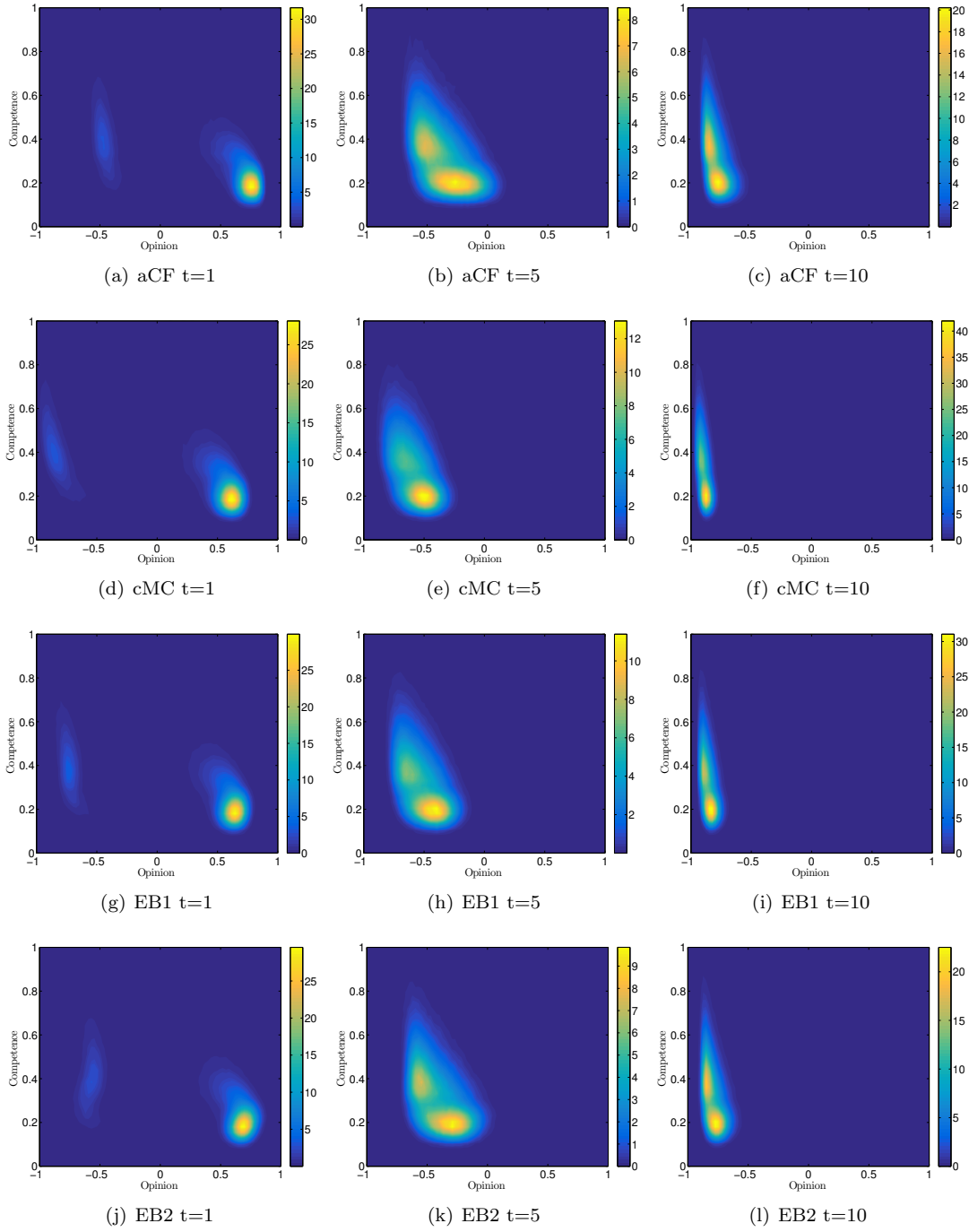


Figure 7: **Test 2.** Kinetic solution for the opinion dynamics with forcing term with rate (5.2) at different time steps. The evolution of the competence variable is given by  $\lambda_B = 10^{-2}$ ,  $\lambda_C(x) = \lambda_C \chi(x \geq x_d)$  where  $\lambda_C = 10^{-2}$  and  $x_d = 0.3$ ,  $\lambda = \lambda_B + \lambda_C$ , and  $\sigma_\kappa^2 = 10^{-2}$ . We considered  $z \sim U([0, 1])$ . We present the behavior of the reference models aCF-cMC and EB1-EB2 under the action of the equality bias functions  $\Phi_1(x), \Phi_2(x)$  for three time steps.

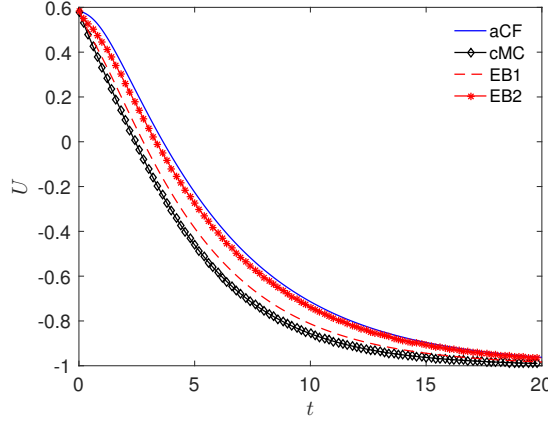


Figure 8: **Test 2.** Convergence of the mean decision in the reference models aCF-cMC and EB1-EB2 under the action of the forcing term with rate (5.2) and  $x_d = 0.3$ . The equality bias function  $\Phi_1(x)$ , which has been sketched in Figure 2, slows down the convergence speed of the collective decision toward  $w_d = -1$ .

where the function  $R(\cdot)$  has been defined in (2.11). We observe how for the choice  $\gamma = 2$  and  $\Delta(x, x_*) = R_{MC}$  the bounded confidence interaction function reproduces the behavior of a maximum competence model.

In the biased case the bounded confidence model becomes

$$P(x, x_*; w, w_*) = \chi(|w - w_*| \leq \gamma \Delta(\Phi(x), \Phi(x_*))). \quad (5.5)$$

We perform the numerical test in the case of absence of driving force, i.e.  $S(\cdot) = 0$ . In Figure 9 we report the initial configuration of the multi-agent system for the test and the asymptotic density function of the opinion variable taking into account the bounded confidence interaction function (5.3) and its biased version (5.5). We chose the parameters  $\lambda_C = \lambda_B = 10^{-3}$ ,  $\lambda = \lambda_B + \lambda_C$  and  $\sigma_\kappa = 10^{-4}$  for the evolution of the competence variable, the function  $\Delta(x - x_*)$  is (5.4) with  $c = 10^2$  and  $\gamma = 1/2$ .

In Figure 10 it is possible to observe how the evolution of opinion and competence deeply changes under the effect of an equality bias function. In particular we considered the equality bias function  $\Phi_2(x)$ ,  $x \in X$  with the choice of parameter introduced in Figure 2. In Figure 9 (right plot) we report the asymptotic marginal density for the opinion variable. The system evolves towards two clusters, characterizing two subpopulations with different decisions driven by the most competent agents. Finally, it is possible to observe how the equality bias drives the system toward a suboptimal collective decision for both populations where the influence of less competent agents become more relevant.

## 6 Conclusion

We introduced and discuss kinetic models of multi-agent systems describing the process of decision making. The models are obtained in the limit of a large number of agents and weight the opinion of each agent through its competence. The binary interaction dynamics involves both the agents' opinion and competence, so that less competent agents can learn during interactions from the more competent ones. This lead to an optimal decision process where the results is a direct consequence of

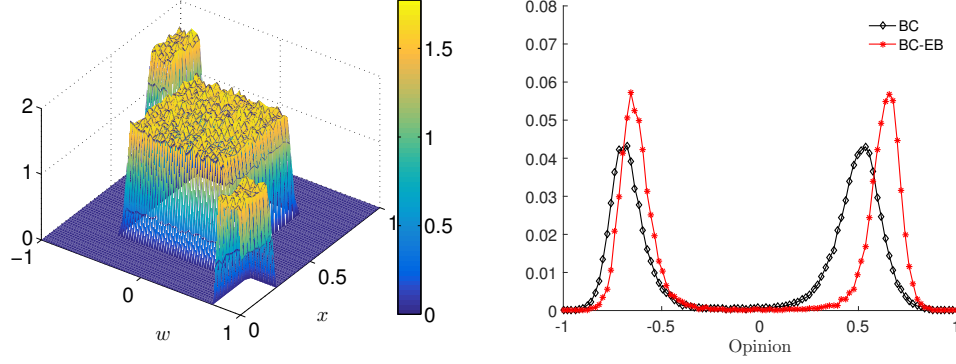


Figure 9: **Test 3.** Left: initial configuration of the multi-agent system, we considered the case with three population. Right: asymptotic marginal density for the opinion variable, we observe how the bounded confidence interaction function introduced in (5.3),  $\gamma = 1/2$ , splits the system in two populations for which the equality bias emerges separately as an action of the equality bias function  $\Phi_2(x)$ .

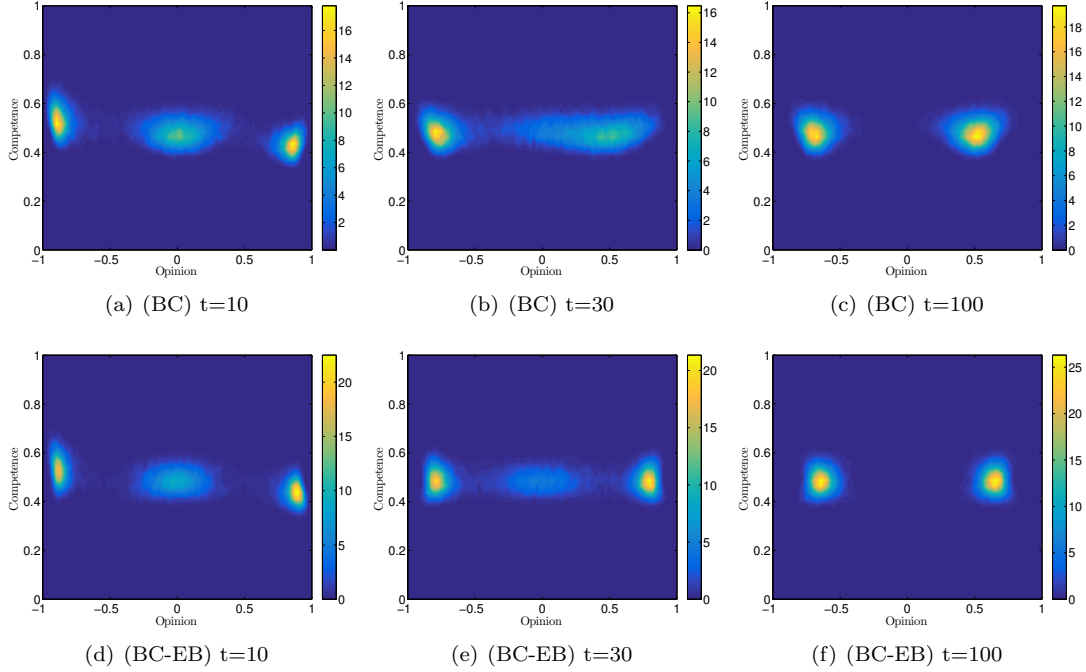


Figure 10: **Test 3.** Evolution of the multi-agent-system in the bounded confidence case (BC), in the first row, and under the action of the equality bias function  $\Phi_2(x)$  (BC-EB), in the second row.

the agents' competence. The introduction of an equality bias in the model is obtained by considering a suitable function with a shape analogous to the one experimentally found in [22]. Numerical results show that the presence of an equality bias leads the group to suboptimal decisions and in some cases to the emergence of the opinion of the less competent agents in the group.

## References

- [1] G. Albi, M. Herty, L. Pareschi. Kinetic description of optimal control problems and applications to opinion consensus. *Communications in Mathematical Sciences*, 13(6): 1407–1429, 2015.
- [2] G. Albi, L. Pareschi, M. Zanella. Boltzmann-type control of opinion consensus through leaders. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 372(2028): 20140138, 2014.
- [3] G. Albi, L. Pareschi, M. Zanella. Opinion dynamics over complex networks: kinetic modeling and numerical methods. *Kinetic and Related Models*, to appear.
- [4] B. Bahrami, K. Olsen, P. E. Latham, A. Roepstorff, G. Rees, C. D. Frith. Optimally interacting minds. *Science*, 329(5995): 1081–1085, 2010.
- [5] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, J. D. Cohen. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice task. *Psychological Review*, 113(4): 700, 2006.
- [6] D. Borra, T. Lorenzi. Asymptotic analysis of continuous opinion dynamics models under bounded confidence. *Communications on Pure and Applied Analysis*, 12(3): 1487–1499, 2013.
- [7] C. Brugna, G. Toscani. Kinetic models of opinion formation in the presence of personal conviction. *Physical Review E*, 92(5): 052818, 2015.
- [8] A. Chakraborti, B. K. Chakrabarti, Statistical mechanics of money: how saving propensity affects its distribution, *The European Physical Journal B-Condensed Matter and Complex Systems* 17(1): 167–170, 2000.
- [9] G. Carbone, I. Giannoccaro. Model of human collective decision-making in complex environments. *The European Physical Journal B*, 88: 339, 2015.
- [10] B. Düring, A. Jüngel, L. Trussardi. A kinetic equation for economic value estimation with irrationality and herding. Preprint arXiv, 2016.
- [11] B. Düring, P. Markowich, J. F. Pietschmann, M. T. Wolfram. Boltzmann and Fokker-Planck equations modelling opinion formation in the presence of strong leaders. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 465(2112): 3687–3708, 2009.
- [12] H. J. Einhorn, R. M. Hogarth. Confidence in judgement: persistence of the illusion of validity. *Psychological Review*, 85(5): 395–416, 1978.
- [13] S. M. Fleming, H. C. Lau. How to measure metacognition. *Frontiers in Human Neuroscience*, 8: 1–9, 2014.
- [14] M. Fornasier, J. Haskovec, G. Toscani. Fluid dynamic description of flocking via Povzner-Boltzmann equation. *Physica D* 240: 21–31, 2011.

- [15] G. Furioli, A. Pulvirenti, E. Terraneo, G. Toscani. The grazing collision limit of the inelastic Kac model around a Lévy-type equilibrium. *SIAM Journal of Mathematical Analysis*, 44: 827–850, 2012.
- [16] S. Galam. Rational group decision making: a random field Ising model at  $T = 0$ . *Physica A. Statistical Mechanics and its Applications*, 238(1–4):66–80, 1997.
- [17] S. Galam, J.-D. Zucker. From individual choice to group decision-making. *Physica A*, 287: 644–659, 2000.
- [18] F. Galton. One vote, one value. *Nature* 75: 414–414, 1907.
- [19] N. Harvey. Confidence in judgement. *Trends in Cognitive Sciences*, 1(2): 78–82, 1997.
- [20] N. Harvey, I. Fischer. Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes* 70(2): 117–133, 1997.
- [21] M. Herty, M. Zanella. Performance bounds for the mean-field limit of constrained dynamics. Preprint, 2015.
- [22] J. Kruger, D. Dunning. Unskilled and unaware of it: how difficulties in recognizing one’s own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology* 77(6): 1121–1134, 1999.
- [23] M. Lallouache, A. S. Chakrabarti, A. Chakraborti, B. K. Chakrabarti. Opinion formation in the kinetic exchange models: spontaneous symmetry breaking transition. *Physical Reviews E* 82(5): 056112, 2010.
- [24] A. Mahmoodi, D. Bang, K. Olsen, Y. A. Zhao, Z. Shi, K. Broberg, S. Safavi, S. Han, M. N. Ahmadabadi, C. D. Frith, A. Roepstorff, G. Rees, B. Bahrami. Equality bias impairs collective decision-making across cultures. *Proceedings of the National Academy of Sciences* 112(12): 3835–3840, 2015.
- [25] S. Motsch, E. Tadmor. Heterophilious dynamics enhances consensus. *SIAM Review* 56(4): 577–621, 2014.
- [26] S. Motsch, E. Tadmor. A new model for self-organized dynamics and its flocking behavior. *Journal of Statistical Physics* 144(5): 923–947, 2011.
- [27] L. Pareschi, G. Russo. An introduction to Monte Carlo methods for the Boltzmann equation. *ESAIM Proceedings*, EDP Sciences. Vol. 10: 35–75, 2001.
- [28] L. Pareschi, G. Toscani. Wealth distribution and collective knowledge: a Boltzmann approach. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 372(2028): 20130396, 2014.
- [29] L. Pareschi, G. Toscani. *Interacting Multiagent Systems. Kinetic Equations and Monte Carlo Methods*. Oxford University Press, 2013.
- [30] I. Poulakakis, L. Scardovi, N. E. Leonard. Coupled stochastic differential equations and collective decision making in the two-alternative forced-choice task. *Proceedings of the 2010 American Control Conference*: 69–74, 2010.
- [31] K. Sznajd-Weron, J. Sznajd. Opinion evolution in closed community. *International Journal of Modern Physics C* 11(6): 1157–1165, 2000.

- [32] G. Toscani. Kinetic models of opinion formation. *Communications in Mathematical Sciences* 4(3): 481–496, 2006.
- [33] K. D. Williams. Ostracism. *Annual Reviews of Psychology* 58: 425–452, 2007.